A new robust approach to volatility forecasting on electricity markets

Luigi Grossi

Department of Economics,
University of Verona

Co-authors:
Lisa Crosato
University of Milano-Bicocca
Fany Nan
Joint Research Centre, European Commission

Objective and basic ideas

- Objective: Robust forecasting of volatility in Electricity Markets.
- Idea: Robust estimators weakly affected by spykes. Focus on the prediction of “ordinary” prices.

GM-Estimator of SETARX processes for the conditional mean equation (Grossi and Nan, 2017). WFS estimator of GARCH processes for the conditional variance equation (Crosato and Grossi, 2017).

Issues:

- Choice of the weighting function for the robust SETARX model;
- Robust tests and criteria for the type and order of models;
- Extension of the classical Forward Search to time dependent data.
- Point forecasting and prediction interval for electricity prices.

SETAR model

Two-regime Self-Exciting Threshold AutoRegressive model SETAR(p,d)

$$y_t = \begin{cases} \mathbf{x}_t \boldsymbol{\beta}_1 + \varepsilon_t, & \text{if} \quad y_{t-d} \leq \gamma \\ \mathbf{x}_t \boldsymbol{\beta}_2 + \varepsilon_t, & \text{if} \quad y_{t-d} > \gamma \end{cases} \tag{1}$$

for $t = 1, ..., N$.

- y_{t-d} threshold variable, $d \geq 1$;
- γ threshold value;
- $\boldsymbol{\beta}_j$ parameter vector for regime $j = 1, 2$;
- $\mathbf{x_t}$, t-th row of the $(N \times p)$ matrix $\mathbf{X}$ comprising p lagged variables of y_t (and a constant, if any);
- $\varepsilon_t \sim iid(0, \sigma_\varepsilon)$.

Robust SETAR model

For a fixed threshold γ the GM estimate of the autoregressive parameters in a two-regime SETAR model can be obtained by the **iterative weighted least squares**:

$$\hat{\beta}_j^{(n+1)} = \left(\mathbf{X}_j'\mathbf{W}^{(n)}\mathbf{X}_j\right)^{-1}\mathbf{X}_j'\mathbf{W}^{(n)}\mathbf{y}_j \ , \tag{2}$$

- $\hat{\beta}_j^{(n+1)}$: GM estimate of parameter vector in regime $j = 1, 2$ after the n-th iteration from an initial estimate $\hat{\beta}_j^{(0)}$,
- $\mathbf{W}^{(n)}$: diagonal weight matrix, whose elements depend on a weighting function $w(\hat{\beta}_j^{(n)}, \hat{\sigma}_{\varepsilon,j}^{(n)})$ bounded between 0 and 1.

The threshold γ can be estimated by minimizing the objective function $\rho(r_t) = \sum_{t=1}^{N} w(\hat{\beta}, \hat{\sigma}_\varepsilon) r_t^2$ over the set Γ.

Polinomial weights

Weights are calculated as

$$w_r(r_t) = \frac{\psi(r_t)}{r_t}$$

for $r_t \neq 0$ and $w_r(0) = 1$.

Following Franses and van Dijk (2000), the ψ function (POL) is:

$$\psi(r_t) = \begin{cases} r_t & \text{if} \quad |r_t| \leq c_1, \\ \operatorname{sgn}(r_t) g(|r_t|) & \text{if} \quad c_1 \leq |r_t| \leq c_2, \\ 0 & \text{if} \quad |r_t| > c_2, \end{cases}$$

where $g(|r_t|)$ is a fifth-order polynomial such that $\psi(r_t)$ is twice continuously differentiable, and c_1 and c_2 are tuning constants.

Eventually, the threshold γ is estimated by minimizing the objective function

$$\rho(r_t) = \sum_{t=1}^{N} w(\hat{\beta}, \hat{\sigma}_\varepsilon)(r_t)^2$$

over the set Γ.

Application to electricity prices. Data Analysis

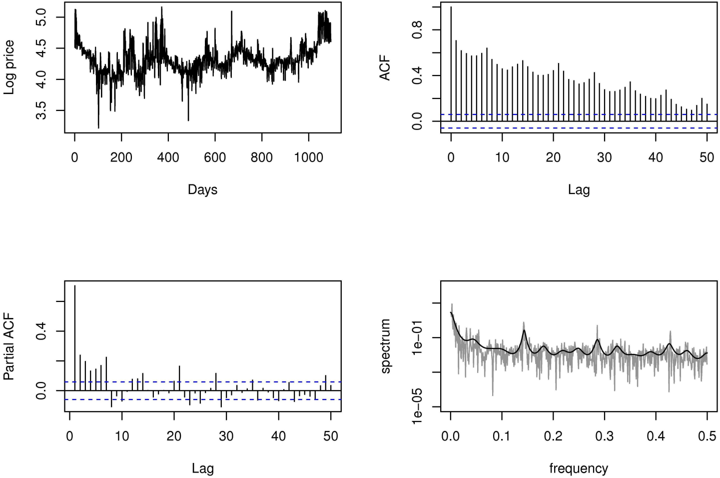
Dataset:

Hourly electricity prices from the Italian Power Exchange (**IPEX**) market.

It is a day-ahead market, as trading typically terminates the day before delivery: prices of 24 hours for next day are fixed once a day by a two-sided auction.

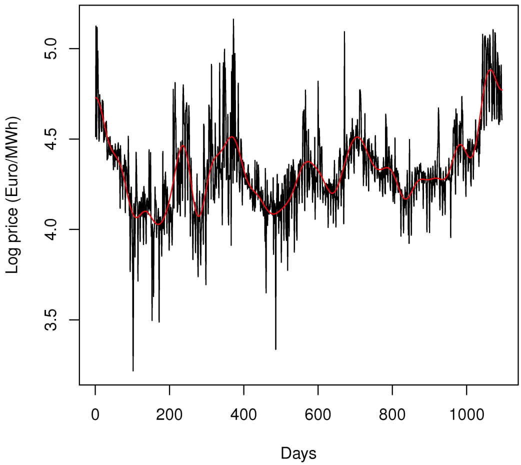
- Every hour is considered separately.
- From 1 January 2013 to 31 December 2015 (26,280 data points).
- Year 2015: 1 step-ahead out-of-sample forecasting.

Log price peak hour 18



Long-run component

- We need to remove the long-run component from prices to obtain a stationary series.
- This component is estimated by Wavelets (Percival and Walden, 2000; Janczura and Weron, 2010)
- Daubechies least asymmetric family; maximum overlap discrete wavelet transform.



Unit root and stationarity tests

Table: *Unit root and stationarity tests applied to original and de-trended time series at 5% (first two columns) and 1% (last two columns) significance levels. Null hypothesis for ADF (Augmented Dickey-Fuller) PP (Phillips-Perron) and ERS (Elliot-Rothenberg-Stock) tests: presence of a unit root. Null hypothesis for KPSS (Kwiatkowski-Phillips-Schmidt-Shin, classic and robust version) tests: stationarity.*

	Number of rejections of the Null Hypothesis			
	Significance level: 0.05		Significance level: 0.01	
Type of Test	**Original**	**De-trended**	**Original**	**De-trended**
ADF	4	24	2	24
PP	24	24	24	24
ERS-DF-GLS	9	13	1	9
ERS-P	5	24	2	24
KPSS	16	0	13	0
Robust KPSS lag7	24	0	24	0
Robust KPSS lag14	24	0	21	0

Non-linearity tests. Non robust

Table: *Number of cases the hypothesis of linearity is rejected*

Hansen's Test (Hansen, 1999)

$d \quad \backslash \quad p$	1	2	3	4	5	6	7
1	18	17	17	17	14	11	10
2	18	10	8	17	17	16	10
3	16	10	8	8	12	13	7
4	18	17	17	11	10	13	15
5	12	13	12	14	13	10	8
6	9	9	11	13	13	15	15
7	22	16	15	14	13	11	12

Non-linearity tests. Robust

Table: *Number of cases the hypothesis of linearity is rejected*

Hung's test (Hung et al., 2009)

$d \setminus p$	1	2	3	4	5	6	7
1	14	14	13	13	13	16	16
2	14	12	12	12	13	16	17
3	14	10	9	10	14	15	11
4	12	15	15	13	14	14	16
5	9	13	12	12	12	11	6
6	13	17	17	17	16	8	10
7	15	15	18	16	16	14	13

Robust model selection

Table: *Robust AIC for different combinations of parameters p (columns) and d (rows). Values in the table are normalized between 0 and 1 for each hour and then averaged over the 24 hours.*

	Robust AIC based on polynomial weighted estimates						
$d \setminus p$	1	2	3	4	5	6	7
1	0.564	0.498	0.475	0.447	0.452	0.323	0.344
2	0.647	0.568	0.584	0.563	0.485	0.410	0.419
3	0.636	0.705	0.564	0.563	0.537	0.422	0.376
4	0.668	0.699	0.682	0.514	0.541	0.398	0.345
5	0.652	0.708	0.630	0.599	0.541	0.406	0.442
6	0.686	0.702	0.669	0.672	0.575	0.451	0.422
7	0.665	0.695	0.634	0.654	0.649	0.526	0.419

Robust model selection

Table: *Robust AIC for different combinations of parameters p (columns) and d (rows). Values in the table are normalized between 0 and 1 for each hour and then averaged over the 24 hours.*

	Robust AIC based on polynomial weighted estimates						
$d \setminus p$	1	2	3	4	5	6	7
1	0.564	0.498	0.475	0.447	0.452	0.323	0.344
2	0.647	0.568	0.584	0.563	0.485	0.410	0.419
3	0.636	0.705	0.564	0.563	0.537	0.422	0.376
4	0.668	0.699	0.682	0.514	0.541	0.398	0.345
5	0.652	0.708	0.630	0.599	0.541	0.406	0.442
6	0.686	0.702	0.669	0.672	0.575	0.451	0.422
7	0.665	0.695	0.634	0.654	0.649	0.526	0.419

Estimation and forecasting design

Estimation:

- SETAR(7,1) models estimated on stationary de-trended series (years 2013-2014);

Exogenous regressors. (Weron, 2014, IJF):

- day-of-the-week dummies, D_k, with $k = 1, \dots, 6$;
- day-ahead predicted demand of electricity (source: Italian Electricity Market Manager, GME).
- day-ahead predicted wind generation (source, TERNA SpA).

Forecasting:

1. out-of-sample forecasting year 2015 of the de-trended prices;
2. predictions of SPOT PRICES (with the long-run component) are avaluated in terms of MSE and MAE;
3. significance of prediction differences are evaluated with the one-tailed Diebold and Mariano test (1995) and Model Confidence Set (MCS) test (Hansen, 2003, 2011) for equal predictive accuracy. Null hypothesis: prediction performance of model A is equal or lower than model B.

Forecasting. Non Robust (LS) vs. Robust (POL) models. Exogenous variables (1)

Table: *SETAR with forecasted demand, dummies and forecasted wind generation: number of cases LS model gives better results than POL model (four seasons and whole year 2015). Comparisons with prediction error statistics (PES) values and p-values for the 1-tailed Diebold-Mariano and MCS tests at 5% significance level, MSE and MAE loss functions.*

	PES Ratios		**D-M test**		**MCS test**	
Period	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**
Jan-Mar	4	2	1	0	0	0
Apr-Jun	10	5	0	0	0	0
Jul-Sep	7	7	0	0	0	0
Oct-Dec	5	6	1	1	0	1
Whole	4	3	0	0	0	0
Totals (120 cases)	25.00%	19.17%	1.67%	0.83%	0.00%	0.83%

Forecasting. Non Robust (LS) vs. Robust (POL) models. Exogenous variables (2)

Table: *SETAR with forecasted demand, dummies and forecasted wind generation: number of cases POL model gives better results than LS model (four seasons and whole year 2015). Comparisons with prediction error statistics (PES) values and p-values for the 1-tailed Diebold-Mariano and MCS tests at* 5% *significance level, MSE and MAE loss functions.*

	PES Ratios		**D-M test**		**MCS test**	
Period	**MSE**	**MAE**	**MSE**	**MAE**	**MSE**	**MAE**
Jan-Mar	20	22	12	13	10	12
Apr-Jun	14	19	5	6	3	7
Jul-Sep	17	17	7	10	3	8
Oct-Dec	19	18	12	12	8	10
Whole	20	21	16	18	15	18
Totals (120 cases)	75.00%	80.83%	43.33%	49.17%	32.50%	45.83%

Robust forward search volatility forecasting

Idea: extend the Forward Search (Atkinson and Riani, 2000; Atkinson, Riani and Cerioli, 2004) to GARCH(1,1) models.

Conflict between FS ranking and time series autocorrelation.

- Previous attempts.
  - State Space representation and Kalman filter estimation of ARMA processes (Riani, 2004);
  - GARCH(1,1) (Grossi, 2004).

THIS WORK: down-weight observations outside the CDS according to their degree of “outlyingness”, then estimate parameters on all observations (Crosato and Grossi, 2017).

Theoretical and empirical framework (1)

GARCH family models (Engle, 1982 and Bollerslev, 1986)

Let ϵ_t denote an observed time series of heteroscedastic residuals. For electricity prices $\epsilon_t = p_t - \hat{p}_t$ where $\hat{p}_t$ is the price fitted by the robust SETAR model.
GARCH(1,1): $\epsilon_t|F_{t-1} \sim N\left(0, \sigma_t^2\right)$ and $\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
with $\alpha_0 > 0, \alpha_1 \geq 0, \beta \geq 0, \alpha_1 + \beta < 1$

- Very large residuals from GARCH models are frequently caused by spikes.
  Additive level outlier: isolated observation affecting the series only at time t, can influence
  - parameter estimates (Zhang and King, 2005; Carnero et al., 2007; Galeano and Tsay, 2010);
  - conditional homoscedasticity tests (Carnero et al., 2007; Grossi and Laurini, 2009)
  - out-of-sample volatility forecasts (Chen and Liu, 1993; Franses and Ghijsels, 1999).

Theoretical and empirical framework (2)

- works on outlier detection in GARCH(1,1)
  - Lagrange Multiplier test (Trivez and Catalan, 2009; Hotta and Tsay, 2012)
  - based on Chen and Liu (1993) recursive method (Franses and Ghijsels, 1999; Charles and Darné, 2005)
  - wavelets (Bilen and Huzurbazar 2002; Grané and Veiga, 2010)

Choice of the initial subset (1)

Window sub-sampling:

group of contiguous observations maintains (part of) the dependence structure of the whole series (Haegerty and Lumley, 2000); the initial subset must be chosen among blocks of contiguous observations, (as in Riani, 2004 and Grossi, 2004)

- T prices
  - divide the T observations into f subsets of contiguous units.
  - Each subset is made up of a sample of $[T/f]$ contiguous observations where $[\cdot]$ indicates the integer part of a number.
- GARCH MLE over $S_h^{(m)}$ with $h = 1, \ldots, f$
- $\widetilde{e}_{t,S_h^{(m)}} = \epsilon_{t,S_h^{(m)}} / s_{t,S_h^{(m)}}$

  standardized residuals for observations in $S_h^{(m)}$

Choice of the initial subset (2)

Least Median of Squares estimator

The initial subset is formed by the observations minimizing the median of squared standardized residuals

$$\min_{h}\left[\widetilde{e}_{[med],S_h^{(m)}}\right]$$

Weighting observations during the search (1)

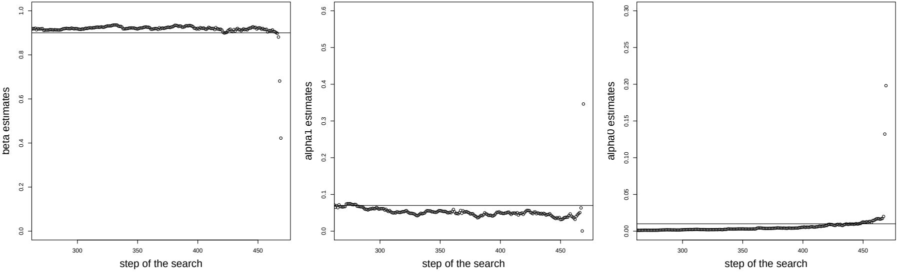
- Let m be the size of the initial Clean Data Set (CDS).
- The remaining $T-m$ observations are ordered according to $\widetilde{e}^2_{t,S^m}$ for $t=1,\dots,T$,
- The WFS now assigns to each observation $\epsilon_{t,j}$ a weight w_j:
  - $w_j = 1, j = 1, \dots, m+1$, so that $\epsilon_{t,1}, \dots, \epsilon_{t,m+1}$ belong to the CDS
  - $w_j = 1 - F_{\chi^2}\left(\widetilde{e}^2_{t,S^m}, 1\right), j = m+2, \dots, T$
    where $F_{res}(\cdot)$ is the squared standardized residuals distribution function.

Weighting observations during the search (2)

- Observations outside the CDS are down-weighted;
- GARCH ML estimates are obtained on the weighted time series, then we move to the next step. Iterative procedure.
- Estimates are based at each step on all observations, no matter the weight

The output of the WFS

Figure: Parameter estimates of a GARCH(1,1) model with $\alpha_0 = 0.01$, $\alpha_1 = 0.07$, $\beta_1 = 0.9$. Contamination: three outliers. Last 200 steps of the WFS.



The output of the WFS

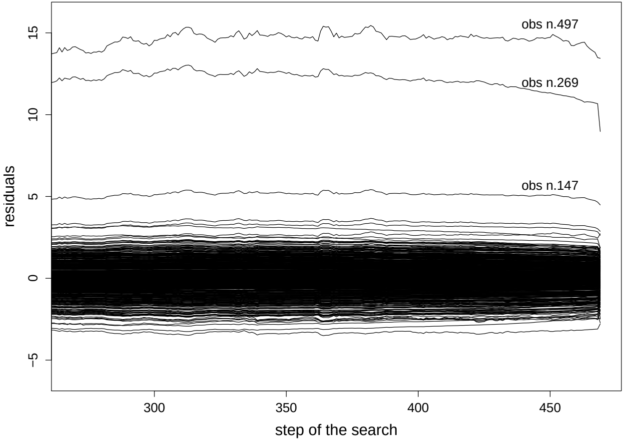
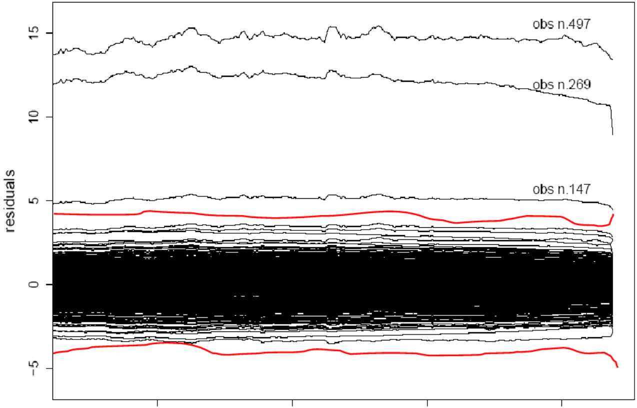


Figure: WFS trajectories of standardized residuals. Last 200 steps of the search.

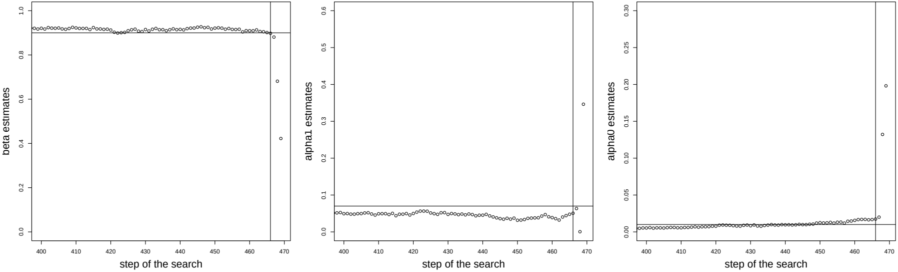
WFS in action. WFS simulated envelopes

Figure: Time series contaminated with three outliers. Three trajectories outside the interval.

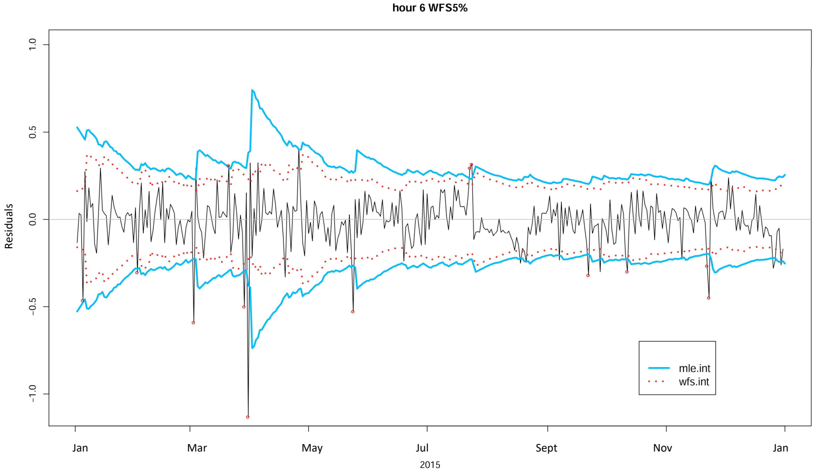


WFS in action

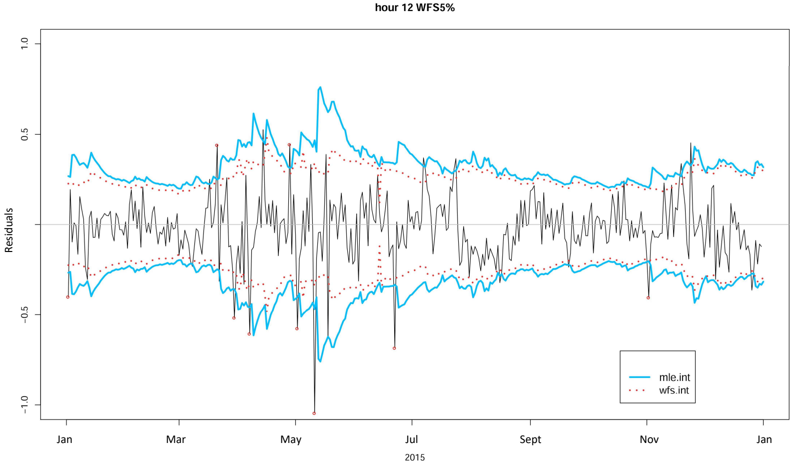
Figure: Time series contaminated with three outliers. Three outliers correctly identified. WFS estimates given by parameters estimated at step T-3.



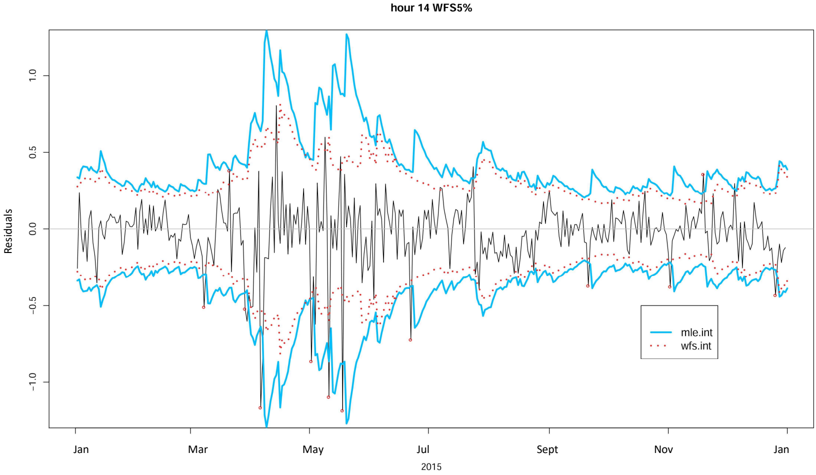
WFS interval on residuals of the conditional mean equation - H6



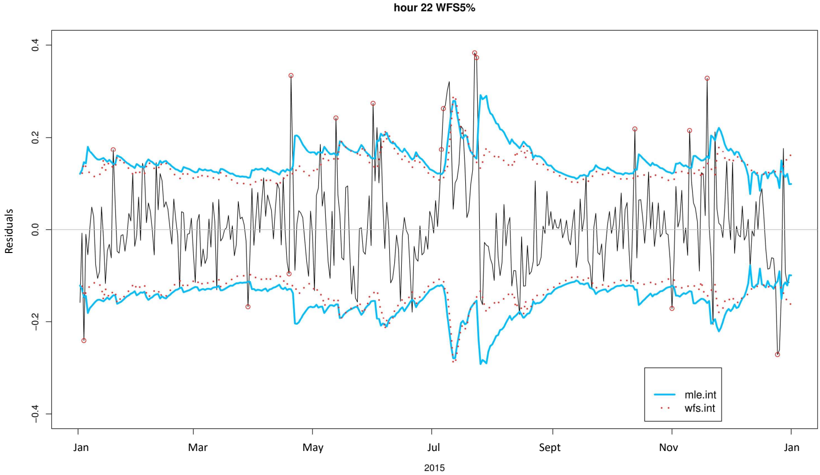
WFS interval on residuals of the conditional mean equation - H12



WFS interval on residuals of the conditional mean equation - H14



WFS interval on residuals of the conditional mean equation - H22



Summary of main results

- An integrated robust procedure for point forecasting and prediction interval of electricity prices.
- Robust SETARX for conditional mean, WFS GARCH for conditional variance; robust test of linearity and robust model selection.
- Good forecasting performance and consistent prediction regions.
- Robust regions applied to electricity prices: more efficient predictions, good outlier detection.

Further research

- WFS algorithm.
  - definition of the initial subset.
  - try different weights?
- Exogenous regression for volatility regions.

References

Atkinson, A.C. and Riani, M., 2000. Robust diagnostic regression analysis, Springer Verlag.
Bilen, C., Huzurbazar, S., 2002. Wavelet-based detection of outliers in time series. Journal of Computational and Graphical Statistics 11, pp. 297-309.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327.
Carnero, M., Peña, D., Ruiz, E., 2007. Effects of outliers on the identification and estimation of GARCH models. Journal of Time Series Analysis 28, pp 471-497.
Chen, C., Liu, L., 1993a. Forecasting time series with outliers. Journal of Forecasting 12, pp. 13-35.
Crosato, L. and Grossi, L., 2017, Correcting outliers in GARCH models: a weighted forward approach, Stat Papers. https://doi.org/10.1007/s00362-017-0903-y
Doornik, J., Ooms, M., 2005. Outlier detection in GARCH models. Tinbergen Institute Discussion Paper.
Franses, P., Ghijsels, H., 1999. Additive outliers, GARCH and forecasting volatility. International Journal of Forecasting 15, pp. 1-9.
Grané, A. and Veiga, H., 2010. Wavelet-based detection of outliers in financial time series. Computational Statistics and Data analysis, 54(11), pp.2580-2593.
Grossi, L., 2004. Analyzing Financial Time Series through Robust Estimators. Studies in Nonlinear Dynamics and Econometrics 8(2), Article 3.
Grossi, L., Laurini, F., 2009. A robust forward weighted Lagrange multiplier test for conditional heteroscedasticity. Computational Statistics and Data Analysis, 53, pp. 2251-2263.
Haegerty, P. J. and T. Lumley, 2000. Window subsampling of estimating functions with application to regression models. Journal of the American Statistical Association, 95, pp. 197-211.
Hotta, L., Tsay, R., 1998. Outliers in GARCH processes. Working paper, IMECC, Brazil and University of Chicago.
Riani, M, 2004. Extensions of the forward search to time series. Studies in Nonlinear Dynamics and Econometrics 8(2), Article 2.
Van Dijk, D., Franses, P., Lucas, A., 1999. Testing for ARCH in the presence of additive outliers. Journal of Applied Econometrics 14, pp. 539-562.
Verhoeven, P., McAleer, M., 2000. Modelling outliers and extreme observations for ARMA-GARCH processes. Working Paper, University of Western Australia.